View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Gene 518 (2013) 17–25

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

An estimation method for a cellular-state-specific gene regulatory network along tree-structured gene expression profiles

Ryo Araki, Shigeto Seno, Yoichi Takenaka, Hideo Matsuda *

Department of Bioinformatic Engineering, Graduate School of Information Science and Technology, Osaka University, 1-5, Yamadaoka, Suita, Osaka 565-0871, Japan

ARTICLE INFO

Available online 21 December 2012

Keywords:
Cellular states
Tree-structured gene expression profiles
Gene regulatory network

ABSTRACT

Background: Identifying the differences between gene regulatory networks under varying biological conditions or external stimuli is an important challenge in systems biology. Several methods have been developed to reverse-engineer a cellular system, called a gene regulatory network, from gene expression profiles in order to understand transcriptomic behavior under various conditions of interest. Conventional methods infer the gene regulatory network independently from each of the multiple gene expression profiles under varying conditions to find the important regulatory relations for understanding cellular behavior. However, the inferred networks with conventional methods include a large number of misleading relations, and the accuracy of the inference is low. This is because conventional methods do not consider other related conditions, and the results of conventional methods include considerable noise due to the limited number of observation points in each expression profile of interest.

Results: We propose a more accurate method for estimating key gene regulatory networks for understanding cellular behavior under various conditions. Our method utilizes multiple gene expression profiles that compose a tree structure under varying conditions. The root represents the original cellular state, and the leaves represent the changed cellular states under various conditions. By using this tree-structured gene expression profiles, our method more powerfully estimates the networks that are key to understanding the cellular behavior of interest under varying conditions.

Conclusion: We confirmed that the proposed method in cell differentiation was more rigorous than the conventional method. The results show that our assumptions as to which relations are unimportant for understanding the differences of cellular states in cell differentiation are appropriate, and that our method can infer more accurately the core networks of the cell types.

© 2012 Elsevier B.V. Open access under CC BY-NC-ND license.

1. Introduction

Systems biology aims to understand cellular processes by using mathematical models (Cantone et al., 2009). One of the important themes in systems biology is to reverse-engineer the dynamics of gene regulatory relations, called gene regulatory networks (GRNs), from gene expression profiles. Several methods have been developed for the mathematical modeling of these dynamics from gene expression profiles. Boolean network (Xiao, 2009), graphical Gaussian model (Grzegorczyk, 2007; Toh and Horimoto, 2002), mutual information model (Margolin et al., 2006), Bayesian network (Heckerman, 1996), and relevance network (Butte and Kohane, 2000; Butte et al., 2003) are widely used.

GRNs have been inferred under various biological conditions, such as stimulated cellular response (Nagashima et al., 2007; Shinozaki et al., 2003) and cell differentiation (Carter et al., 2004; Siersbak and Mandrup, 2011; Siersbak et al., 2012; Tokuzawa et al., 2010; Zhang et al., 2011). The aim of this research is to identify the key gene regulatory relations in a given cellular state; this is because the key relations in a state help us to understand the differences between various cellular states. Cellular states are described by GRNs in this area of research; different states have different GRNs, and the states are changed by the conditions. To find the key relations in cellular states, conventionally, the GRN of the cellular state has been inferred from a single gene expression profile under the condition of interest.

It is not an easy task for conventional methods to identify the key gene regulatory relations in a cellular state. This is because conventional methods run into two problems. The first is that conventional methods infer the GRN of a cellular state from a single gene expression profile under the condition of interest; these are called *exact samples* (Hu, 1994; Hu et al., 2000). With the conventional methods, the inferred GRN includes relations in common with other states

Abbreviations: GRN, gene regulatory network; GEO, Gene Expression Omnibus; GEP, gene expression profile; ROC, receiver operating characteristic; MSC, mesenchymal stem cell; RN, relevance network; GGM, graphical Gaussian modeling.

* Corresponding author.
E-mail address: matsuda@ist.osaka-u.ac.jp (H. Matsuda).

0378-1119 © 2012 Elsevier B.V. Open access under CC BY-NC-ND license.
http://dx.doi.org/10.1016/j.gene.2012.11.090

and relations specific to the given state. Another problem is that for the GRNs inferred from the expression profiles that are derived from public databases such as Gene Expression Omnibus (GEO), more than 90% of profiles have fewer than 10 observation points. The result is that the inferences produced by the conventional methods include a large number of candidate relations that are common to other states or unique to the given state. Due to above two problems, the inferred GRNs of conventional methods have low accuracy.

In this paper, we propose a more accurate method of inferring the key GRN in a cellular state from multiple gene expression profiles under varying biological conditions. It should be noted that conventional methods aim to infer a GRN under a condition of interest, and they are not intended to compare multiple gene networks, that is, to infer each GRN independently from *exact samples*. Our method is different from these conventional methods in that it utilizes the exact profile along with profiles from other conditions, called *relevant samples* (Hu, 1994; Hu et al., 2000), for inferring the key GRN in a cellular state. We assume that the key GRN in a cellular state is a relation specific to that state, and that removing common relations from the GRN of the state enables a more accurate inference of the key GRN. We also use multiple gene expression profiles that compose tree structure, and in this tree, both *exact sample* and *relevant samples* are included. The root of the tree corresponds to an original state of a target state and the comparison states, and the leaves represent the target state and the comparison states. The profiles from the root to the target leaf are merged and used as a profile of the target state, and the profiles from the root to the comparison leaves are merged and used as profiles of the comparison state. These merged profiles decrease the candidate relations in the inferred GRNs, and the specific relations are extracted from the GRN of the target by comparing them with the GRNs of the comparisons. Therefore, our method can estimate the key GRN of the target state with higher accuracy.

2. Method

In this study, we defined cellular-states-specific GRNs as a set of gene regulatory relations that determine the differences of cellular states. It is difficult for conventional methods to extract the relations specific to a cellular state because, as mentioned, the small number of observation points in the profiles leads the inferred GRN to include a large number of candidates, and not to consider other cellular states lead to include relations common and specific to the other states. Thus, our method considered the other states related to target state and decreased the candidate relations by applying tree-structured gene expression profiles. Below is the notation we will use for the explanation of our methods.

2.1. Notation

$S=\{S_1,\ldots,S_m\}$: A set of cellular states.
S_i: An original cellular state of S_i.
$G=\{g_1,\ldots,g_n\}$: A set of genes g.
$GEPs=\{GEP_1,\ldots,GEP_m\}$: A set of gene expression profiles of S.
$GEP_i[G,T_i]$: A gene expression profile including observation points T_i and a set of gene G.
$T_i=\{t_1^i,\ldots,t_j^i\}$:A set of observation points t of GEP_i.

We explain our tree-structured profiles with this notation. To simplify the explanation we will describe our method with three gene expression profiles. This example is easily applied to more gene expression profiles. One profile is observed from S_0 to S_1, a second is observed from S_1 to S_2 and another is from S_1 to S_3 (see Fig. 1). Let S_2 be a target cellular state, then the conventional methods infer a target GRN from the profile that is observed from S_1 to S_2. The results include key interactions for understanding the behavior of a target state; however, the inferred GRN has too many candidate relations to determine the key relations of this behavior because the profile has only a few observation points and the conventional inference of the GRNs does not consider other cells S_3. For understanding the differences in the behaviors, it is necessary to utilize other profiles and to compare the GRNs in order to detect the differences between the target cellular state and the other cellular states. Thus, our methods use other profiles both for decreasing the candidate relations in the inferred GRNs and for removing the relations common to S_2 and S_3.

For extracting a GRN specific to a cell, we assumed that there are two kinds of common relations in the GRNs. One is the gene relations common to all cellular states, such as the relations maintaining cellular states. A conceptual view of this assumption is shown in Fig. 2. Another assumption is that there are relations common to a particular series of cellular states, such as relations inherited from an original state S_0. We assume that we cannot distinguish the behavior unique to a target state from these relations (see Fig. 3). And common relations to a particular series of states cannot be estimated from profiles with few observation points, because the relations are quiet.

We propose two methods following from these assumptions for a more accurate estimation of cellular-states-specific relations.

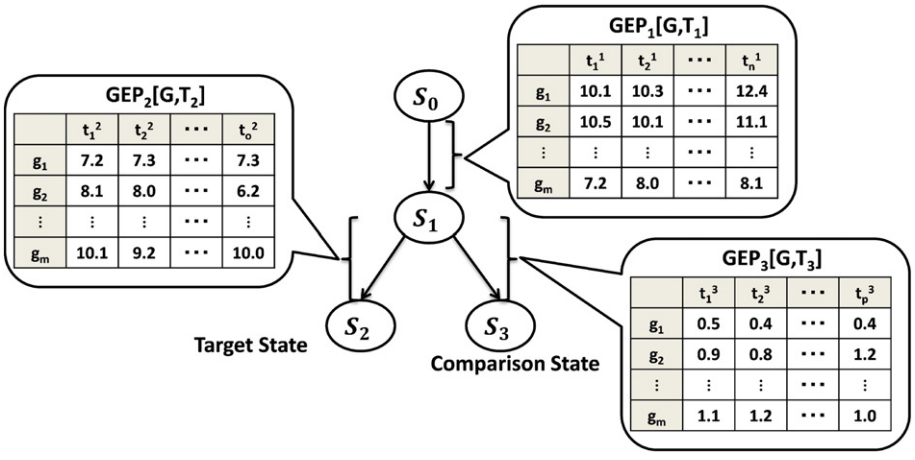

Fig. 1. Tree-structured gene expression profiles. This figure shows an example of tree-structured gene expression profile composed of 3 profiles. The circled S corresponds to cellular states. One profile observes the transitions from the root S_0 to S_1, another profile is from S_1 to S_2, and the third is from S_1 to S_3. Cellular states are changed from S_0 through S_1 to S_2 and S_3.

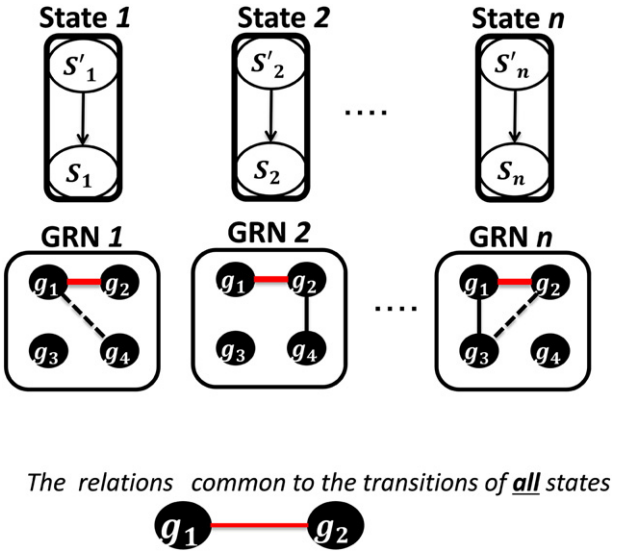

Fig. 2. Conceptual view of relations common to all cellular states. This figure shows the assumption that all cellular state transitions have the same gene relations, e.g., a positive relation between g_1 and g_2 in this figure. S represents a cellular state and S' corresponds to an original cellular state of S. Here, g corresponds to a node of the gene regulatory network. The solid edges between nodes are activating relations and dashed edges are suppressing interactions. All cellular state transitions from S'_i to S_i, are different; however, all changes have the same activation or suppression interactions between g_1 and g_2.

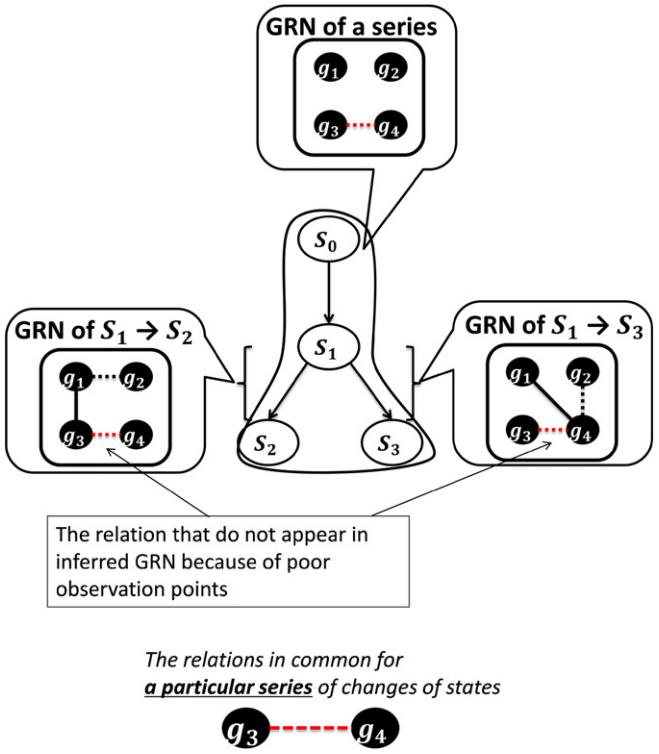

Fig. 3. Conceptual view of relations common in a particular series. A conceptual view of this assumption is shown in this figure. In this assumption, a particular series of cellular states has the same interactions between genes, such as between g_3 and g_4. In this figure, S_0, S_1, S_2, and S_3 are surrounded by a line, and this is a series of S_0 posterity.

2.2. Purpose and overview

Our purpose is to enhance the accuracy of the inference of cellular-states-specific GRNs for understanding the behavior of cellular states. To achieve this purpose, we applied tree-structured gene expression profiles to decrease the candidate relations in the inferred GRNs and to compare the GRNs. We explain these as described above, with tree-structured profiles composed of three gene expression profiles. In the following, the GRNs are described by a weighted adjacency matrix.

A conceptual view of step flow is shown in Fig. 4.

2.2.1. Overview of Method 1

Method 1 follows the assumption that all cellular states have common relations. This method is composed of four steps. In Step 1, the input data are normalized with an initial expression value of each gene. This process is important for the merging step (Step 2) below because if the gene expression profiles have significantly different values, it would be too difficult to estimate the gene relations. Step 2 is a merging step. This step is the main step of our method to solve the problem of low accuracy of the inference GRNs caused by too few observation points in the gene expression profiles. In this step, we merge the input data to increase the observation points of the target profile and of the comparisons. As shown in Fig. 4, the merged profiles are those observed from S_0 to S_1 and from S_1 to S_2. Another merged profile, from S_0 to S_2, is used as the profile of the target state S_2. The profiles observed from S_0 to S_1 and from S_1 to S_3 are merged and used as a profile of a comparison state. In Step 3, the GRNs are estimated from the profiles created in Step 2. The estimation method is the same as that for the conventional methods. Let AM be a weighted adjacency matrix of GRN, and its elements $AM[i, j]$ represent gene regulatory relations between g_i and g_j. The gene regulatory relations are represented as an adjacency matrix $AM[i, j]$. The matrix has the same number of rows and columns, and the number equates to the number of genes in G. The element of $AM[i, j]$ indicates whether there is a regulatory relation between g_i and g_j; the value of elements is -1, 0, or 1. When the element of $AM[i, j]$ is -1, there is a negative relation between g_i and g_j, the value of 0 corresponds to no relation and 1 represents positive relations between the genes. The value of the element is determined by whether the criteria exceed a threshold. Several criteria have been proposed along conventional mathematical modeling methods of inferring GRN, such as correlation, partial correlation, mutual information, a parameter of a differential equation, and Bayesian probability. Convenient criteria were used that consider the trade-off between computational time and accuracy. Finally, we compared the GRNs inferred in Step 3, and the relations in common between the target and comparison states were removed from the GRN of the target. The residues are output as a target-state-specific GRN.

2.2.2. Overview of Method 2

Method 2 extends Method 1 with the assumption that there are common relations in a given series of cellular states. The series is composed of multiple cellular states included in the path from the root cellular state to target and comparison cellular states. Method 2 also has four steps, which are the same as for Method 1 except for Step 2. Step 2 of this method is the merging step, which extends the merging step of Method 1. In this step, all expression profile are merged (from S_0 to S_1, from S_1 to S_2 and from S_1 to S_3), and this merged profile is used as the profile of a series of target and comparison states. The expression profiles of the target and comparison states are merged in the same way as in Method 1. Those relations that are held in common by the target GRN, the comparison GRN, and the GRN inferred from the profile of the series, are removed from the target GRN in Step 4. Then, the remaining relations in the target GRN are output as the target-state-specific GRN.

2.3. Algorithms

As a constraint, input data *GEPs* compose a tree structure, and all input profiles have the same set of genes G.

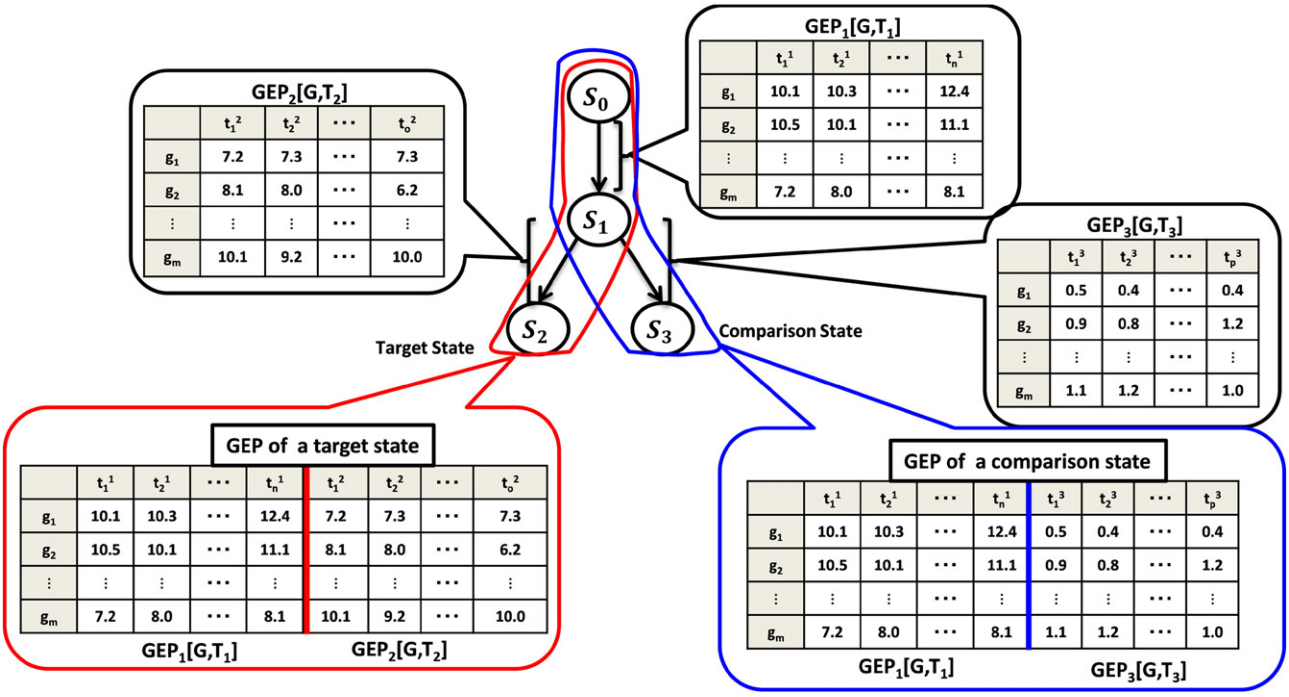

Fig. 4. Conceptual view of the merging step in Method 1. In this figure, we show a conceptual view of the merging step in Method 1. All *GEP*s have the same gene set, and GEP_1 is observed from S_0 to S_1, GEP_2 from S_1 to S_2, and GEP_3 from S_1 to S_3. The profile created by merging GEP_1 and GEP_2, surrounded by a red line, is used as a profile of the target state S_2. The profile created by merging GEP_1 and GEP_3, surrounded by a blue line, is used as a profile of the comparison state S_3.

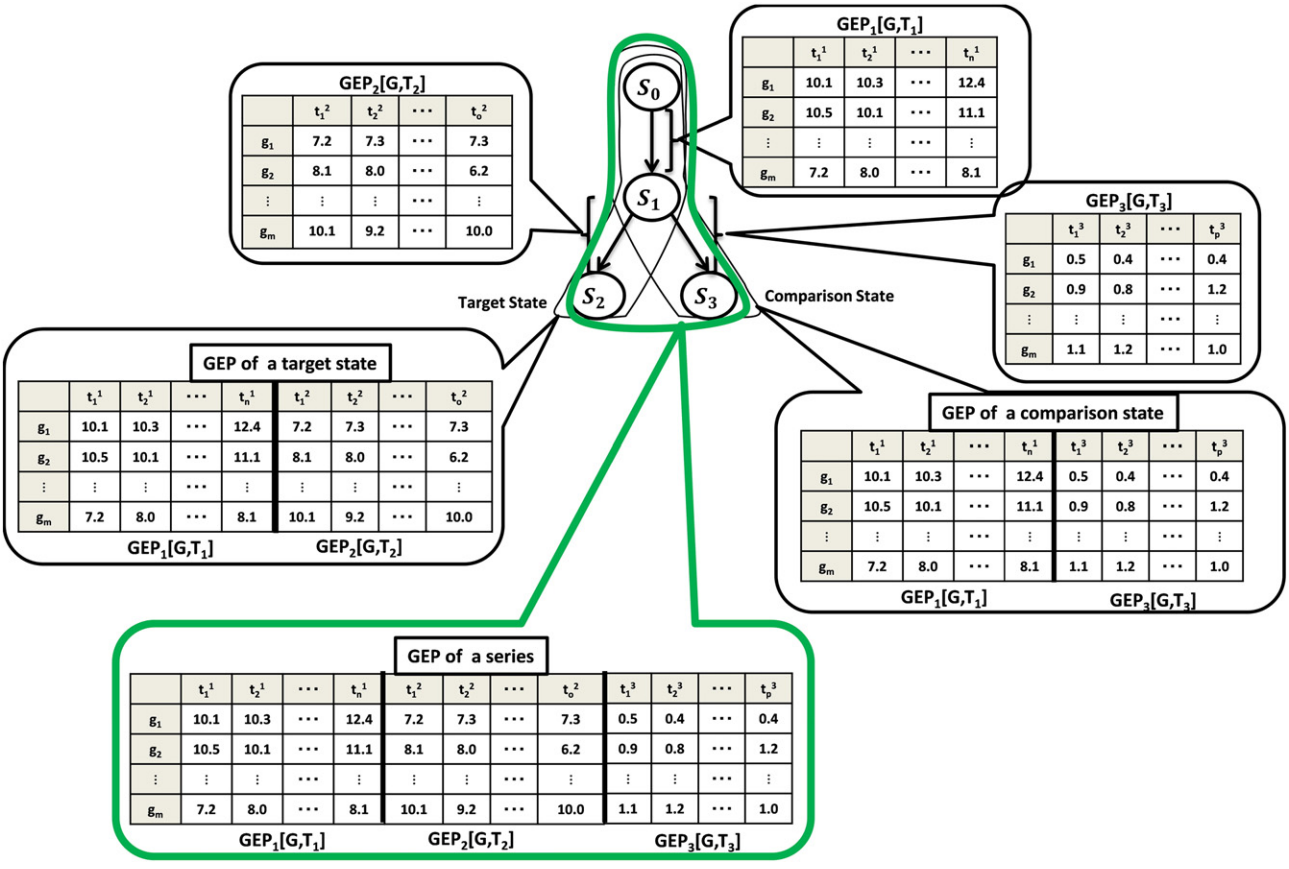

Fig. 5. Conceptual view of the merging step in Method 2. This figure shows the merging step in Method 2. The merged profile of the target state S_2 and the comparison state S_3 are the same as in Method 1. This step is different from the step in Method 1 in that all the *GEP*s are merged as a single profile of a series of S_0, S_1, S_2, and S_3.

Input:

$GEPs = \{GEP_1, \ldots, GEP_n\}$: a set of multiple gene expression profiles composing a tree structure;
TC: threshold of criteria.

Output:

A_{tar}: a target-state-specific GRN matrix;
A_{com}: a comparison specific GRN matrix.

Step 1: *Normalizing input gene expression profiles*
In this step, all gene expressions of each of the observation points were normalized. In this study, all gene expression values were divided with each initial expression value as a representative value. This normalization was for merging profiles that have very different expression values. Without this step, estimation of GRNs from merged profiles having large differences of value would not be efficient.
Step 2: *Merging normalized profiles*
In this step, the normalized profiles in Step 1 are merged to increase the observation points of the profiles of the target state and the comparison states. The output of this step is normalized and merged profiles, and decrease the number of candidate regulatory relations in the inferred GRNs. This step is a little different in Method 1 than in Method 2, so we will explain this step separately for each method.

Step 2 in Method 1
This step follows the assumptions that all transitions of cellular states have the same gene relations (see Figs. 2 and 4, and the left-hand side of Fig. 6). All input expression profiles compose tree structure, which includes the target state. Every profile belongs to either the path from the root to the target state or from the root to the comparison states. Some profiles belong to both paths (to target and to comparison), and such profiles are merged several times. This merging process is executed for every profile.
Variable:

$MG_{tar}[G,T_{tar}]$: MG_{tar} corresponds to the merged target profile including the observation points *T* and genes *G*;
$MG_{com}[G,T_{com}]$: MG_{com} is merged comparison profile including the observation points *T* and genes *G*.

First, we checked which paths the profile belonged, and this checking was done by a person. Secondly, if the profile was part of the path to the target state, then the profile was connected to the end of MG_{tar}. When the profile belonged to the path to the comparison states, connect the profile to the end of MG_{com}. As mentioned above, some profiles belonged to both paths, and such profiles were connected to the end of both MG_{tar} and MG_{com}. These processes ere repeated for all the input profiles.
Step 2 in Method 2
In this step, we follow the assumption that there are common relations in a given series of states (see Figs. 3 and 5 and the right-hand side of Fig. 6). This step is very similar to Step 2 in Method 1, in that it produces a merged profile of the target state MG_{tar} and a merged profile of the comparison states MG_{com}. In addition to these merged profiles, this step, also produced a merged profile of a series that includes the target and the comparison states. The profile of the series is produced from all input profiles because the input profiles compose the tree, including the target and the comparison states, and this input tree profile is defined as a series of the target and the comparison.

Variable:

$MG_{tar}[G,T_{tar}]$; MG_{tar} corresponds to the merged target profile including observation points *T* and genes *G*;
$MG_{com}[G,T_{com}]$: MG_{com} is the merged comparison profile including observation points *T* and genes *G*;
$MG_{all}[G,T_{all}]$: MG_{all} represents the merged profile of the series including observation points *T* and genes *G*.

The process of creating M_{tar} and MG_{com} is the same as the process in Method 1. To make the merged profile of the series, all input profiles are, one after another, connected to the end of MG_{all}.

Step 3: *Estimating the GRNs*
In this step, we estimated by using conventional criteria the gene regulatory relations from the several merged profiles that were output by Step 2. We also described the inferred GRN with a weighted adjacency matrix facilitates removing the relations held in common. After this process, relations were represented by three values; positive interaction, 1; negative interaction, −1; and no relationship, 0. The process of this step is explained below in pseudo-code.

```
for MG_i ⊆ (MG_tar[G, T_tar], MG_com[G, T_com] and MG_all[G, T_all]) do
  for all g_j ⊆ G do
    for all g_k ⊆ G do
      CM_i[j, k] ← calculate criteria (MG_i[g_j, T_i], MG_i[g_k, T_i])
      RM_i[j, k] ← calculate reliability of criteria (MG_i[g_j, T_i], MG_i[g_k, T_i])
      if CM_i[j, k] > TC then
        AM_i[j, k] ← 1
      else if CM_i[j, k] < TC then
        AM_i[j, k] ← (−1)
      else then
        AM_i[j, k] ← 0
      end if
      if RM_i[j, k] < calculate reliability (TC,num(T_i)) then
        AM_i[j, k] ← 0
      end if
    end for
  end for
end for
```

Step 4: *Removing common relations*
In this step, relations held in common to the target GRN and the comparison GRN are removed from both the target and the comparison GRNs, and this step outputs a target-state-specific GRN and a comparison-state-specific GRN. The relations that are held in common are found by comparing the elements of the adjacency matrix of target and comparison GRNs. When the values corresponding to the gene relations are also in the target and comparison matrices, the values are changed to 0 to indicate there is no relationship. This process is described below in pseudo-code.

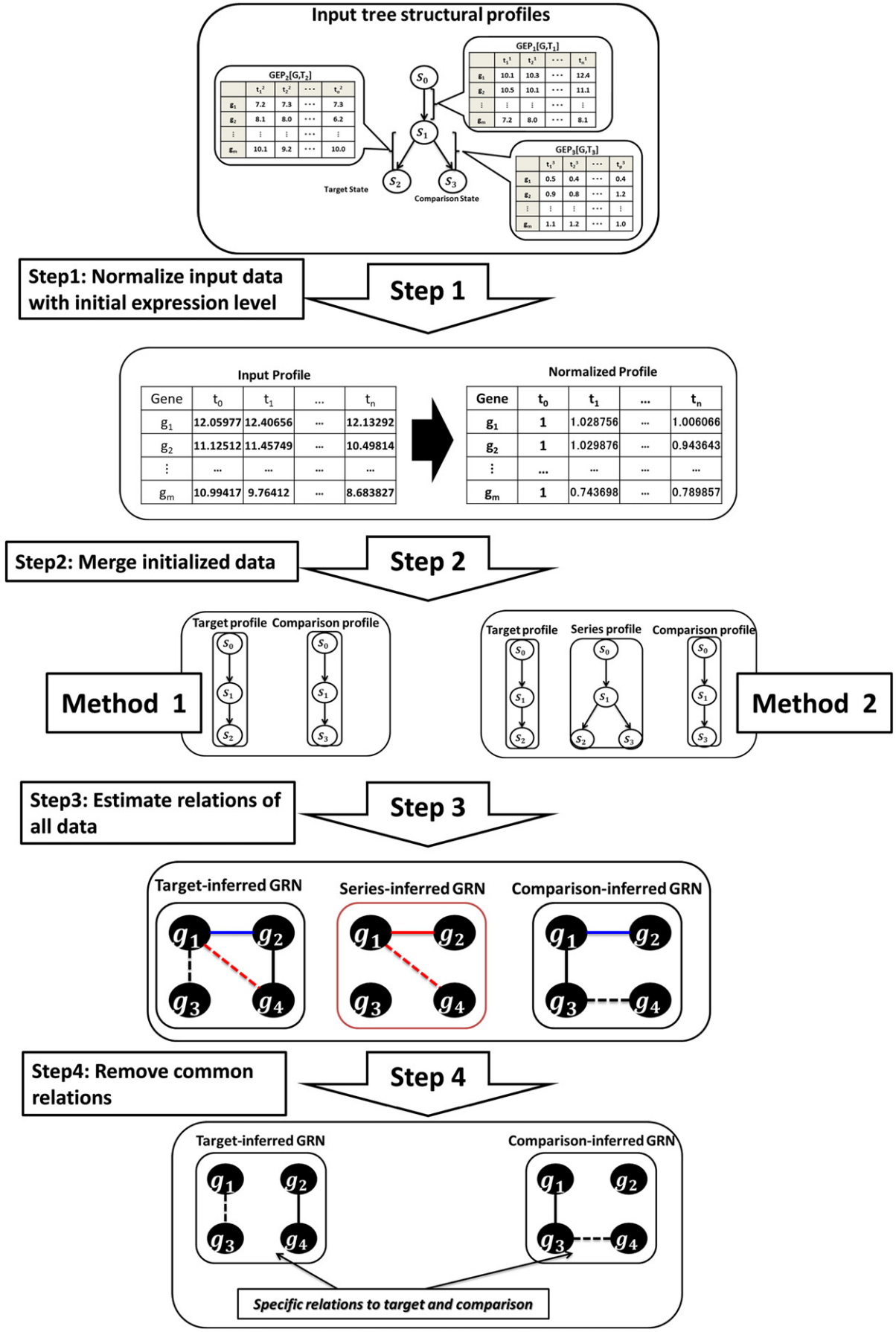

Fig. 6. Conceptual view of a flow chart of the steps. In this figure, the labeled arrows indicate the executable steps.

01:**for all** $AM_i \subseteq (AM_{tar}, AM_{com}, AM_{all})$ **do**

02: $A_i \leftarrow AM_i$

03:**end for**

04:**for all** $j \subseteq G$

05: **for all** $k \subseteq G$

06: **if** $AM_{tar}[j,k] == AM_{com}[j,k]$ **then**

07: $A_{tar}[j,k] \leftarrow 0$

08: $A_{com}[j,k] \leftarrow 0$

09: **end if**

10: **if** $AM_{tar}[j,k] == AM_{all}[j,k]$ **then**

11: $A_{tar}[j,k] \leftarrow 0$

12: **if** $AM_{com}[j,k] == AM_{all}[j,k]$ **then**

13: $A_{com}[j,k] \leftarrow 0$

14: **end for**

15:**end for**

16:**return** A_{tar} and A_{com}

3. Results and discussion

3.1. Results

We evaluated the proposed method by comparing it with conventional methods about the accuracy of the inferred cellular-states-specific GRNs. We used the receiver operating characteristic (ROC) curve for comparing the accuracy, and we extract a certain threshold of criteria for comparison accuracy. In our experiment, we used two gene expression profiles that observe cell differentiation. We inferred the gene relations that may decide the destination of the cell differentiation or cell type's difference. The input data were the observed differentiation into adipocyte or osteoblast of mesenchymal stem cell (MSC) of *Mus musculus* (the house mouse). They are time series gene expression profiles. One data set of MSC changing into adipocyte has 7 observation points (2, 4, 6, 8, 10, 12, and 14 days), and the other data set of MSC changing into osteoblast has 15 observation points (1, 6, 12, 18, 24, 30, 36, and 42 h; and 2, 4, 6, 8, 10, 12 and 14 days). These data are available from Genome Network Platform (http://genomenetwork.nig.ac.jp). We extracted 14 genes that were already known to compose the core network of adipocyte or osteoblast. Already known core networks of adipocyte and osteoblast are shown in Fig. 7 We used Pearson's correlation as the criteria of relations for estimating the GRN, which is called a relevance network (RN) (Butte and Kohane, 2000; Butte et al., 2003). The correlation gives a value between $+1$ and -1, and the relations have strong regulatory relationships as the value that is almost $+1$ or -1. Conventional methods infer the GRN with RN from only one profile. Estimation methods were implemented with R 2.13.0. We defined the already known relations as true positive relations and the others as true negative relations (see Appendix A for the technical details of the experiment).

First, we drew the ROC curve and compared the conventional and proposed methods by changing the criteria threshold. The horizontal axis corresponds to false positive rate (equal to 1-specificity), and the vertical axis is sensitivity. The ROC curve is shown in Fig. 8. In Fig. 8, the black solid line corresponds to the conventional method, the blue line is proposed Method 1, and the red line is Method 2. The left-hand side shows the results of the adipocyte-specific GRN estimation, and the right-hand side shows the results for osteoblast. The conventional method was no more accurate than random estimation, shown as a dashed black line on the ROC curve. On the other hand, each proposed method was more accurate than both random and conventional estimations. Proposed Method 2 was more accurate than the Method 1 in the results of adipocyte and osteoblast.

Second, for the criteria threshold that correlation was more than 0.8 or less than -0.8, we compared the accuracy. The results are shown in Fig. 9, and we compared, in this figure, false negative rate and positive rate of the conventional method and our method. The left-hand side corresponds to the adipocyte result and the right-hand side corresponds to the osteoblast result. The false negative rates of the three methods are almost the same on both sides. However, false positive rate of the proposed methods are less than that of the conventional method, and the false positive rate of the proposed Method 2 is less than that of Method 1. From this result, our methods are more accurate than the conventional method.

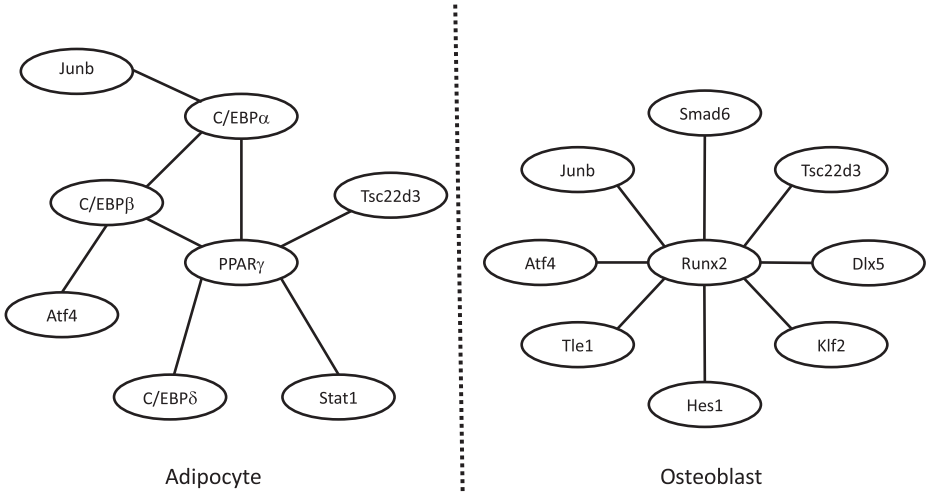

Fig. 7. Already known core networks. Core network relations that are already known; the left-hand side of the figure corresponds to adipocyte, and the right-hand side corresponds to osteoblast.

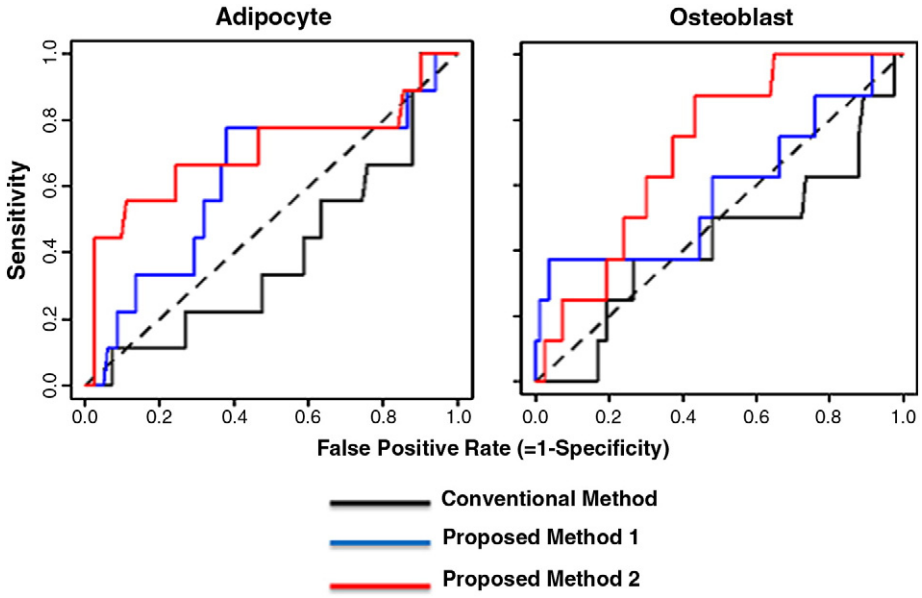

Fig. 8. The ROC curve of adipocyte- and osteoblast-specific GRN estimations. The dashed black line represents the random network, the solid black line is the result of the conventional method, and the blue and red lines are the results of our methods. The vertical axis represents the sensitivity of the estimated GRN, and the horizontal axis is false positive rate in the results. The left-hand side of the figure displays the results for adipocyte, and the right-hand side displays those for under osteoblast.

3.2. Discussion

The results show that the proposed methods are appropriate in this condition, MSC differentiation, and Method 2 is more appropriate than Method 1. This is because, the inference of the GRN from one expression profile includes relations held in common and relations specific to each cell type, adipocyte and osteoblast, thus the conventional method cannot infer only specific relations as we expected. Our methods extract specific relations by comparing with the other cell types. Method 2 removes relations held in common to the series of adipocyte and osteoblast, and these relations don't compose the core network of these cells. Thus, our proposed methods are more rigorous in that it extracts the cellular-states-specific core GRN.

As a future work, inferred networks in these experiments are undirected networks although gene regulatory relations have direction, so it might be better that we use other estimation methods with direction, such as GGM. And in merging steps, we merged profiles that have different time intervals, so we should consider the differences of time interval between profiles.

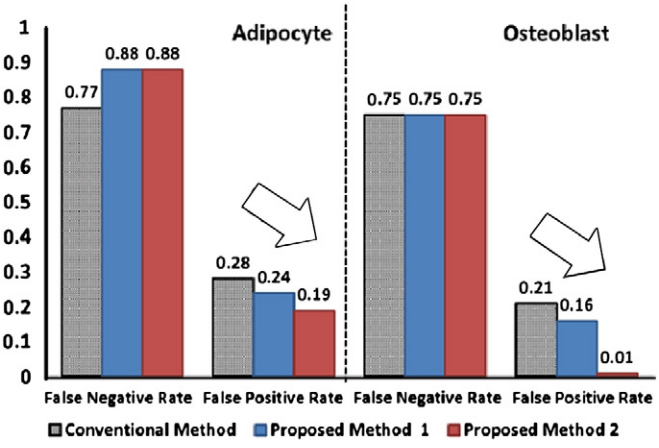

Fig. 9. Comparison of the accuracy using fixed threshold criteria. The left-hand side of this figure shows the result of the estimation of the adipocyte-specific GRN, and the false negative rate and false positive rate are compared. The gray bar corresponds to conventional method, and the blue and red bars correspond to proposed Methods 1 and 2, respectively. The right-hand side shows the same information, but for the osteoblast.

4. Conclusions

In this paper, we present a novel estimation method for inferring the cellular-states-specific gene regulatory network. In our methods, we also propose two assumptions to gene–gene relations held in common. We confirmed that the proposed methods are able to estimate cellular-specific relations more accurately than the conventional method.

Author's contributions

RA implemented the algorithm and performed the analyses. RA, SS, YT, and HM conceived and designed the experiments and wrote the paper.

Acknowledgments

This work was partially supported by a Grant-in-Aid for Scientific Research (22680023 and 22310125) from the Japan Society for the Promotion of Science (JSPS), and by the HPCI STRATEGIC PROGRAM Computational Life Science and Application in Drug Discovery and Medical Development from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT).

Appendix A. Technical details of the experiment

The data utilized in our experiment is expression microarray data of mouse ST2 differentiating into adipogenesis/osteoblastogenesis. The microarray platform is Affymetrix GeneChip Mouse Genome 420 2.0 array. Each time-course data was background-subtracted and normalized with the robust multi-array analysis (RMA) using affy package from the Bioconductor version1.8.1.

The tree structure is composed of three nodes; MSC, adipocyte and osteoblast. We considered an integrated profile as a MSC series profile.

Output of our experiment is an undirected adjacency square matrix of 14 rows and columns. The rows and columns corresponded to the 14 genes shown in Fig. 7.

References

Butte, A., Kohane, I., 2000. Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. Pac. Symp. Biocomput.

Butte, A.S., Kohane, I.S., Parmigiani, G., Garrett, E.S., Irizarry, R.A., Zeger, S.L., 2003. Relevance networks: a first step towards finding genetic regulatory networks within microarray data. The Analysis of Gene Expression Data, pp. 428–446.

Cantone, I., et al., 2009. A yeast synthetic network for in vivo assessment of reverse-engineering and modeling approaches. Cell 137, 172–181.

Carter, S.L., Brechbühler, C.M., Griffin, M., Bond, A.T., 2004. Gene co-expression network topology provides a framework for molecular characterization of cellular state. Bioinformatics 20 (14), 2242–2250.

Grzegorczyk, M., 2007. Extracting protein regulatory networks with graphical models. Proteomics 51–59.

Heckerman, D., 1996. A tutorial on learning with Bayesian networks. Innovations in Bayesian Networks, pp. 33–82.

Hu, F., 1994. Relevance Weighted Smoothing and A New Bootstrap Method. Department of Statistics, University of British Columbia.

Hu, F., Rosenberger, W.F., Zidek, J.V., 2000. Relevance weighted likelihood for dependent data. Metrika 51, 223–243.

Margolin, A.A., et al., 2006. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinforma. 7 (Suppl 1), S7.

Nagashima, T., et al., 2007. Quantitative transcriptional control of ErbB receptor signaling undergoes graded to biphasic response for cell differentiation. J. Biol. Chem. 282 (6), 4045–4056.

Shinozaki, K., Yamaguchi-Shinozaki, K., Seki, M., 2003. Regulatory network of gene expression in the drought and cold stress responses. Curr. Opin. Plant Biol. 6 (5), 410–417.

Siersbak, R., Mandrup, S., 2011. Transcriptional networks controlling adipocyte differentiation. Cold Spring Harb. Symp. Quant. Biol.

Siersbak, R., Nielsen, R., Mandrup, S., 2012. Transcriptional networks and chromatin remodeling controlling adipogenesis. Trends Endocrinol. Metab. 23 (2), 56–64.

Toh, H., Horimoto, K., 2002. Inference of a genetic network by a combined approach of cluster analysis and graphical Gaussian modeling. Bioinformatics 18 (2), 287–297.

Tokuzawa, Y., et al., 2010. Id4, a new candidate gene for senile osteoporosis, acts as a molecular switch promoting osteoblast differentiation. PLoS Genet. 6 (7), e1001019.

Xiao, Y., 2009. A tutorial on analysis and simulation of Boolean gene regulatory network models. Curr. Genom. 10 (7), 511–525.

Zhang, L., Zheng, Y., Li, D., Zhong, Y., 2011. Self-organizing map of gene regulatory networks for cell phenotypes during reprogramming. Comput. Biol. Chem. 35 (4), 211–217.